

**TRENCH ISOLATION FOR INTEGRATED CIRCUIT**Field of the Invention

5           The invention relates generally to integrated circuit design and process technology. In particular, the invention pertains to a trench isolation structure and a method for filling isolation trenches.

Background of the Invention

10           In silicon integrated circuit technology, a large number of isolated devices are fabricated on the same semiconductor substrate. Subsequently these devices are interconnected along specific paths to create the desired circuit configuration. In the continuing trend toward higher densities, devices are placed closer together with smaller spaces between them. Device isolation, which is critical to proper circuit operation, has  
15           become progressively more challenging.

          Over the last few decades a variety of isolation technologies have been developed to address the requirements of various integrated circuit types. In general, the various isolation technologies are different with respect to characteristics such as minimum isolation spacing, surface planarity, process complexity and defect density  
20           generated during isolation processing. Moreover, it is common to make trade-offs among these characteristics when developing an isolation process for a particular integrated circuit application.

          In MOS (metal-oxide-semiconductor) technology, LOCOS isolation (LOCAl Oxidation of Silicon) has been the most widely used isolation technology. LOCOS  
25           isolation involves thermal oxidation of a silicon substrate through a mask. The resultant field oxide is generally grown thick enough to avoid cross-talk between adjacent devices, but not so thick as to cause step coverage problems. The great popularity of LOCOS isolation technology can be attributed also to its inherent simplicity in MOS process integration, its cost effectiveness and its adaptability.

30           In spite of its success, several limitations of LOCOS technology have driven the development of alternative isolation structures. A well-known limitation in LOCOS

isolation is oxide undergrowth at the edge of the hard mask (typically made of silicon nitride) that protects the active regions of the substrate during oxidation. This so-called "bird's beak" poses a limitation to device density since it widens the isolation region, thereby reducing resolution, while causing stress within the substrate. Another problem associated with the LOCOS process is its non-planarity. For submicron devices, non-planarity becomes an important issue, often posing problems for photolithography and subsequent layer conformality.

Trench isolation technology has been developed, in part, to overcome the aforementioned limitations of LOCOS isolation for submicron devices. Refilled trench structures comprise a recess formed in the silicon substrate, which is filled with a dielectric material. Such structures are fabricated by first forming submicron-sized trenches in the silicon substrate, usually by a dry anisotropic etching process through openings in a photoresist overlayer. Typically the resulting trenches display a steep sidewall profile. The trenches are refilled with a dielectric material, such as silicon dioxide ( $\text{SiO}_2$ ), usually deposited by a chemical vapor deposition (CVD) or spin-on glass (SOG) process. Then the surface is smoothed by an etchback or polishing process so that the trench is completely filled with the dielectric material, and its top surface is level with the silicon substrate. After a successful etchback, the top surface has good planarity and is at the same level as the surrounding substrate. The resulting structure serves to electrically isolate adjacent devices.

Refilled trench isolation can take a variety of forms depending upon the specific application. It is generally categorized in terms of the trench dimensions: shallow trenches ( $<1\ \mu\text{m}$  deep), moderate depth trenches ( $1\text{-}3\ \mu\text{m}$  deep), and deep, narrow trenches ( $>3\ \mu\text{m}$  deep,  $<2\ \mu\text{m}$  wide). Shallow Trench Isolation (STI), for example, is used primarily for isolating devices of the same type and is often considered an alternative to LOCOS isolation. Furthermore, shallow trench isolation has the advantages of eliminating the bird's beak of LOCOS and providing a high degree of surface planarity.

The basic trench isolation process is, however, subject to drawbacks, including void formation in the trench during dielectric refill. Voids result when the refilling dielectric material forms a constriction near the top of the trench before it is completely

filled, as shown in Figure 1A. Such voids compromise device isolation as well as the overall structural integrity of the integrated circuit. Unfortunately, preventing void formation during trench refill often places minimum size constraints on the trenches themselves, which can limit device packing density or the effectiveness of the device isolation. A key parameter in measuring the effectiveness of device isolation is the field threshold voltage, that is, the voltage necessary to create a parasitic current linking adjacent isolated devices. The field threshold voltage is influenced by a number of physical and material properties, such as trench width, dielectric constant of the trench filling material, substrate doping, field implant dose and substrate bias.

Generally, void formation has been mitigated by decreasing trench depth and/or tapering trench sidewalls so that the openings are wider at the top than at the bottom, as shown in Figure 1B. A principal trade off in decreasing the trench depth is reducing the effectiveness of the device isolation, while the larger top openings of trenches with tapering sidewalls use up additional and valuable integrated circuit real estate.

Accordingly, it is desirable to develop a trench isolation process that overcomes the problem of void formation while providing effective device isolation.

#### Summary of the Invention

In accordance with one aspect of the present invention, a method of fabricating trench isolation structures between integrated electrical devices in a semiconductor substrate is provided. The method includes filling the trenches with insulating material by atomic layer deposition, comprising a plurality of primary cycles. Each primary cycle comprises, in sequence, introducing a first vapor-phase reactant to the substrate, thereby forming no more than about one monolayer of a first reactant species, conforming at least to surfaces of the trenches, removing excess first vapor-phase reactant and byproduct from the reaction chamber, introducing a second vapor-phase reactant to the substrate, thereby reacting with the first reactant species conforming at least to the surfaces of the trenches, and removing excess second vapor-phase reactant and byproduct from the reaction chamber.

In accordance with another aspect of the invention, a method is provided for forming a trench isolation structure in an integrated circuit. The method includes

forming a trench in a semiconductor substrate and filling the trench with an insulating material comprising a mixture of silicon oxide and aluminum oxide.

5 In accordance with another aspect of the invention, a method is provided for fabricating a trench-fill structure in an integrated circuit. The method includes forming a trench within a structural material in a partially fabricated integrated circuit. A first filler material is deposited by a plurality of primary atomic layer deposition cycles, each of which deposits no more than about one monolayer of the first filler material. A second filler material is deposited by a plurality of secondary atomic layer deposition cycles, each of which deposits no more than about one monolayer of the second filler material. The primary and secondary cycles are performed a number of times to fill the trench. The ratio of primary cycles to secondary cycles is selected to form a composite filler material with a coefficient of thermal expansion (CTE) within about 20% of the structural material's CTE.

15 In accordance with another aspect of the present invention, an integrated circuit is described that comprises a semiconductor substrate with a plurality of neighboring active device regions, openings embedded within the semiconductor substrate between the neighboring active device regions and an insulating material that fills the openings conformally, the insulating material comprising a mixture of silica and mullite. In the illustrated embodiment, the ratio of mullite to silica in the insulating material is preferably between about 25 wt % mullite/75 wt % silica to about 50 wt % mullite/50 wt % silica, resulting in a coefficient of thermal expansion between about  $2.0 \times 10^{-6}$  /K and  $3.0 \times 10^{-6}$  /K.

25 In accordance with another aspect of the invention, a trench isolation structure in an integrated circuit is described that comprises an opening embedded within a semiconductor substrate between neighboring devices and an insulating material that fills the opening conformally and that has a linear coefficient of thermal expansion within about 20% of the linear coefficient of thermal expansion of the semiconductor substrate.

### Brief Description of the Drawings

These and further aspects of the invention will be readily apparent to those skilled in the art from the following description and the attached drawings, wherein:

5 Figure 1A is a schematic cross-section of a partially fabricated integrated circuit, illustrating a filled trench with a void embedded within a semiconductor substrate.

Figure 1B is a schematic cross-section of a partially fabricated integrated circuit, illustrating a filled trench with tapered sidewalls.

Figure 2 is a schematic, perspective view of a single-substrate reaction chamber configured for atomic layer deposition.

10 Figure 3 is an exemplary atomic layer deposition gas flow diagram in accordance with a preferred embodiment, wherein silicon is deposited and then oxidized.

15 Figure 4 is an exemplary atomic layer deposition gas flow diagram in accordance with a preferred embodiment, wherein silicon is deposited and oxidized and then aluminum is deposited and then oxidized.

Figure 5 is a schematic cross-section of a partially fabricated integrated circuit, illustrating isolation structures formed between active devices.

Figure 6 is a schematic cross-section of a partially fabricated integrated circuit, illustrating a partially filled trench.

20 Figure 7 is a schematic cross-section of a partially fabricated integrated circuit, illustrating the partially filled trench of Figure 6 at a later stage of deposition.

Figure 8 is a schematic cross-section of a partially fabricated integrated circuit, illustrating a filled trench.

25 Figure 9 is a schematic cross-section of a partially fabricated integrated circuit, illustrating the filled trench of Figure 8 after planarization.

Figure 10 is an equilibrium phase diagram for silicon oxide (silica) and aluminum oxide (alumina).

Figure 11 is a plot of coefficient of thermal expansion as a function of weight percent  $\text{Al}_2\text{O}_3$  in an  $\text{Al}_2\text{O}_3/\text{SiO}_2$  mixture.

30

### Detailed Description of the Preferred Embodiment

Although described in the context of trench isolation structures in integrated circuits, the skilled artisan will readily find application for the principles disclosed herein in a number of other contexts. The processes and structures of the preferred  
5       embodiments have particular utility where extremely thin layers are desired, particularly within openings that are narrow and deep.

As noted in the Background section above, the trend in integrated circuit technology is toward increased miniaturization of devices. As devices are made  
10       smaller, other integrated circuit structures, such as isolation trenches, also become smaller and are therefore more and more difficult to fill by conventional means. For example, Figure 1A illustrates schematically a trench isolation structure comprising a narrow trench 120 in a semiconductor substrate 100 that has been filled with a dielectric  
15       filler layer 140 by conventional chemical vapor deposition (CVD). As shown, the deposited filler layer 140 pinches off at the opening of the trench 120 before the trench is filled completely. The premature pinching off of the deposited layer 140 results in a void within the trench. Such voids are unacceptable for device isolation within the substrate.

A common solution to the problem of premature pinching off is to taper the trench sidewalls. Figure 1B is a schematic illustration of an exemplary approach to  
20       prevent pinching off during conventional chemical vapor deposition (CVD). The sidewalls of the trench are tapered so that the deposited film 150 adequately fills the bottom of the trench 170 before the top is sealed. One disadvantage of the tapered trench, however, is that the trench 170 takes up extra space at the surface of the substrate 100 because of the wide opening, which ultimately limits circuit density.

Accordingly, prevailing development trends in integrated circuit technology  
25       have placed ever increasing demands on thin film deposition techniques. The continuous decrease in the size of devices and increase in design complexity require nearly perfect conformality and accurate film thickness control. It is an aspect of the present invention to provide a technique that addresses these requirements.

30       A thin film deposition technique that may meet these requirements is atomic layer deposition (ALD), known also as atomic layer epitaxy (ALE). ALD is a chemical

vapor phase, thin film deposition technique wherein film growth proceeds by exposing the wafer surface to pulses of precursor gases. Individual precursors are pulsed onto the surface of the wafer alternately and sequentially and removed between the pulses, such as by evacuation or by purging with an inert gas. In one pulse, no more than about one monolayer of the precursor is absorbed onto the surface through chemisorption in a self-limiting reaction that stops when the surface is saturated. The gas in the next pulse either reacts directly with the monolayer or at least removes the self-limiting surface termination, thus allowing further reaction with a subsequent pulse. Films are built up by growing no more than about one monolayer at a time in a cycle of short bursts of gases, and thickness is controlled by the number of reaction cycles.

The principles of ALD are described by T. Suntola, for example, in the Handbook of Crystal Growth 3, Thin Films and Epitaxy, Part B: Growth Mechanisms and Dynamics, Chapter 14, Atomic Layer Epitaxy, pp. 601-663, Elsevier Science B.V. 1994 and in U.S. Patent No. 4,058,430, the disclosures of both of which are incorporated by reference herein in their entirety. Theoretically, ALD films can have perfect uniformity and conformality over a large area, regardless of non-uniformities in reactant concentrations and temperatures across the area. Accordingly, a layer can be deposited to conform to and fill very narrow and deep trenches, even with large substrates, such as 300-mm silicon wafers.

Moreover, the inventors have realized that ALD also affords greater compositional flexibility. This realization has been applied to another source of stress and/or voids in semiconductor processing. In the course of fabricating an integrated circuit, the substrate is subjected to multiple thermal cycles, in which various materials undergo different thermal expansions. Disparate materials can thereby withdraw from one another (causing voids) or compress one another (causing stress). This is particularly true of trenches formed in a structural layer and filled with a different filler material. For example, in a conventional shallow trench isolation context, silicon oxide has a different coefficient of thermal expansion (CTE) than the silicon substrate that surrounds it. In part because of the excellent control enabled by ALD, materials can be selected to minimize differential thermal expansion while achieving the traditional goals of device isolation.

Figure 2 is a schematic representation of an ALD reactor 180 including a reaction chamber 182 configured to optimize deposition by ALD. This figure is a general illustration of a reaction chamber. It is not meant to be represent any specific reaction chamber, nor is it meant to include all necessary components of a reaction chamber. The process recipes provided herewith can be usefully employed in a Pulsar™ 2000 ALCVD™ reactor, commercially available from ASM Microchemistry Oy of Espoo, Finland. The skilled artisan can readily adjust process parameters for other reactors in view of the disclosure herein.

Gases flow into the reaction chamber 182 through a plurality of gas inlets 184. Vapor-phase reactants, such as precursor gases and oxidizing gases, and carrier gases communicate with the inlets 184. Advantageously, the inlets 184 are arranged to keep reactant gases separate until opening into the chamber 182, thereby avoiding undesirable CVD-type reactions and consequent particulate formation. Note that the terms “vapor-phase reactants” and “reactant gases” are meant to encompass even reactants that are considered to be in a liquid or solid phase at the reaction conditions, as long as they have a vapor pressure high enough to saturate the substrate surface in a reasonably short cycle.

Carrier gas can be any inactive gas suitable for conveying vapor-phase reactant gases through the chamber 182 and also can also be employed to purge the chamber between reactant gas pulses. The reactant gases pass over the substrate 186 where atomic layer deposition and oxidation can occur. Generally, though not necessarily, heaters (not shown) are also provided to heat the substrate to within an ALD temperature window, which depends upon the particular process used. Unused precursor gases, oxidizing gases, reaction products and carrier gases leave the chamber 182 through the gas outlet 188.

#### **Preferred ALD Film Deposition Processes**

It is difficult to deposit conformal layers to fill very narrow openings, for example narrower than 100 nm, with conventional fabrication methods, particularly where the opening has an aspect ratio greater than about 3:1 (depth:width). The preferred embodiments, however, employ atomic layer deposition (ALD), which forms thin films monolayer by monolayer. Advantageously, these atomically thin monolayers



result in conformal deposition of filler material within very narrow openings or trenches.

In general, a structural material includes a trench to be filled with a different material. In the illustrated embodiment, the structural material comprises a semiconductor substrate, particularly a silicon wafer, having a trench defined therein to be filled with an insulating material. The substrate surface may be the top portion of an intrinsically doped silicon wafer or an epitaxial silicon layer. In other arrangements, the substrate can comprise alternative materials, such as III-V or other semiconductors. Trench-fill isolation can also be employed for silicon-on-insulator (SOI) technology. In still other arrangements, the structural material can comprise a thick, planarized insulating layer in which trenches or contact vias are formed, to be filled with a conductive material.

Surfaces of the trench within the structural material (substrate) are prepared to leave a surface termination that readily reacts with the first reactant in the ALD process. In the illustrated embodiment, wherein a dielectric layer is to be formed over a single-crystal silicon wafer, the bare silicon surface is preferably terminated with hydroxy (OH) tails. As will be appreciated by the skilled artisan, such a surface termination can be readily obtained by an *ex situ* HF dip, followed by exposure to the clean room atmosphere. In other arrangements, the substrate can be prepared by an *in situ* H<sub>2</sub>O vapor or alcohol/carboxylic acid vapor treatment. Exemplary alcohols include methanol, ethanol and propanol. An exemplary carboxylic acid is acetic acid.

In accordance with the preferred embodiment, a substrate with trenches is surface-treated for atomic layer deposition, either before or after being loaded into the reactor chamber. After the substrate is loaded, an inactive carrier gas flow purges the chamber to remove any atmospheric contaminants. Preferably, the carrier gas is nitrogen, helium or argon and flows from the inlet to the outlet. In the illustrated embodiment, the carrier gas is nitrogen. An exemplary gas flow rate is about 0.5 - 1.0 slm (standard liters per minute).

Temperature and pressure process parameters can be modified for individual film materials and for desired film characteristics. For an ALD silicon oxide process using 3-aminopropyltrimethoxysilane as a silicon precursor, for example, it is preferable

to ramp the process temperature to a range of from about 150°C to 400°C, and more preferably, from about 200°C to 300°C. The pressure range is preferably from about 3 Torr to 50 Torr, and, more preferably, from about 5 Torr to 20 Torr.

5 In an alternative embodiment of the present invention, ALD of silicon oxide can take place at even lower temperatures. Using remote-plasma excited oxygen and/or nitrogen sources, even room temperature processing is possible. As will be appreciated by the skilled artisan, plasma processes generally employ lower pressures, such as in the millitorr (mTorr) range. Accordingly, any combination of precursor gases and carrier gas may be directed separately to the plasma generator which couples microwave  
10 energy from a magnetron to the flowing gases, thus generating excited species for an alternative process.

After the substrate reaches the desired reaction temperature and the chamber is set to the desired pressure level, precursor gases are released into the continuous carrier gas flow according to directions programmed into a central controller. The precursor  
15 gases are injected as pulses into the carrier gas flow and are thus distributed into the process chamber. These precursor gases flow through the process chamber, react with the substrate, and are drawn toward the outlet or exhaust port. Unreacted precursor gases and any gaseous reaction byproducts are purged from the chamber after the reaction through the outlet port by the continuously-flowing carrier gas.

20 A first precursor gas or vapor-phase reactant is pulsed into the chamber by injection into a carrier gas flow. A first chemical species from the precursor gas is adsorbed onto the prepared substrate. In the illustrated embodiments, wherein the desired films include silicon oxide, the first species comprises silicon and at least one other ligand. The result is a self-terminating monolayer of the silicon species.  
25 Preferably, the silicon source gas comprises a gas conducive to self-limiting chemisorption, such as silicon halides, halosilanes and organic silicon compounds. Examples include silicon tetrachloride ( $\text{SiCl}_4$ ), dichlorosilane (DCS or  $\text{SiH}_2\text{Cl}_2$ ), trichlorosilane (TCS or  $\text{SiHCl}_3$ ), species of 3-aminoalkyltrialkoxysilane, such as 3-aminopropyltriethoxysilane ( $\text{NH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{-Si(O-CH}_2\text{CH}_3)_3$  or AMTES) or 3-aminopropyltrimethoxysilane ( $\text{NH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{-Si(O-CH}_3)_3$  or AMTMS), or other  
30 halosilane silicon sources. When the precursor is a liquid, such as TCS, the delivery

system can include a bubbler and a gas line for bubbling N<sub>2</sub> through the liquid to transport silicon-containing gas molecules more effectively to the reaction chamber in gaseous form.

In an exemplary embodiment, the substrate comprises a silicon wafer having isolation trenches formed therein, and the substrate is maintained at a temperature between about 150°C and 400°C, more preferably between about 200°C and 300°C, while reactor pressure is maintained between about 5 Torr and 20 Torr. The skilled artisan will readily appreciate that reaction conditions can be varied for different reactors and different chemistries.

The silicon source gas comprises 3-aminopropyltriethoxysilane and is maintained for between about 0.5 seconds and 3.0 seconds under the preferred temperature and pressure conditions, and more preferably for between about 1.0 seconds and 2.0 seconds. No more than about monolayer of silicon chemisorbs onto the silicon substrate surface and terminates with organic or halide ligands. The surface termination of the monolayer inhibits further reaction with the silicon source gas.

After the silicon source pulse is turned off, carrier gas is allowed to flow for between about 0.5 seconds and 4.0 seconds, more preferably between about 1.0 seconds and 2.0 seconds, to purge excess source gas and reaction byproducts from the chamber prior to the next reactant pulse. In other arrangements, it will be understood that the chamber can be evacuated to remove the reactant species and byproducts.

Next, a second vapor-phase reactant is provided to the substrate by injection into the carrier gas flow. In the illustrated embodiment, the second reactant comprises an oxygen or oxidant source that reacts with the ligand termination, either organic or halide, of the adsorbed silicon monolayer, leaving oxygen atoms in place of the ligands.

Oxidizing gases such as ozone are preferred for silicon layers that terminate in organic ligands. The ozone is preferably injected into the carrier gas flow at a concentration of about 1 - 30 volume percent, more preferably about 2 - 25 volume percent. Under the preferred temperature and pressure conditions, the ozone pulse is maintained for between about 0.5 seconds and 5.0 seconds, more preferably between about 2.0 seconds and 4.0 seconds.

Oxidizing gases such as water vapor are preferred for silicon layers that terminate in halide ligands. The water vapor is preferably injected into the carrier gas flow. Under the preferred temperature and pressure conditions, the water vapor pulse is maintained for between about 0.2 seconds and 5.0 seconds. Preferably, stoichiometric or near-stoichiometric SiO<sub>2</sub> is produced.

After the oxygen source pulse is turned off, carrier gas is allowed to flow for about 0.5 seconds to 4.0 seconds, preferably at least about 1.0 second, to purge the oxidizing source gas and reaction byproducts from the chamber prior to the next reactant pulse. In other arrangements, it will be understood that the chamber can be evacuated to remove the reactant species and byproducts.

Table I below provides an exemplary process recipe for one cycle of silicon oxide deposition.

**TABLE I: Exemplary SiO<sub>2</sub> Deposition Cycle**

Phase	Carrier Flow (slm)	Reactant	Temperature (°C)	Pressure (Torr)	Time (sec)
silicon	0.8	3-aminopropyltriethoxysilane	300	10	1.5
purge	0.8	--	300	10	1.5
oxidant	0.8	ozone	300	10	3.0
purge	0.8	--	300	10	1.5

In accordance with the principles of ALD, a second pulse of the silicon source gas is then injected into the carrier gas flow. The pulse is stopped, and the silicon source gas is removed from the chamber, followed by a second oxygen source gas pulse, which is then, in turn, stopped and removed from the chamber. These pulses are then continually alternated until the silicon oxide layer attains the desired thickness. Alternatively, the cycle of Table I can be mixed with cycles depositing other materials, as discussed below with respect to the embodiment of Figure 4 and exemplary Table II (Al<sub>2</sub>O<sub>3</sub>). For convenience, Table I is labeled a "SiO<sub>2</sub>" deposition cycle, though when mixed with other cycles it may form other materials. Table I is also referred to herein as representing a "primary cycle." Use of the term "primary" is for labeling purposes only, and does not imply relative importance of the material formed thereby nor the sequence in which different ALD cycles are employed.

Figure 3 is a gas flow diagram in accordance with the preferred embodiments, illustrating the first four cycles 301a - 301d in an exemplary ALD process, wherein silicon dioxide is deposited. There is a continuous flow of a carrier gas 300 throughout the process. As shown, a first pulse or spike 302a of the silicon source gas is provided to form the first self-terminated silicon monolayer or less. After a first purge step 303, during which carrier gas continues to flow until the silicon source gas has been removed from the chamber, a first oxygen source gas pulse or spike 304a is provided. After a second purge 303, a second silicon source gas pulse 302b is provided, followed by a second oxygen source gas 304b, a third silicon source gas pulse 302c, a third oxygen source gas pulse 304c, etc., in alternating pulses separated by purge steps 303. Note that Figure 3 is schematic only and is not drawn to scale.

Actually, in practice, the ALD process forms a full monolayer only after a plurality of cycles. Theoretically the reactants chemisorb at each available site on the exposed layer of the workpiece, but the physical size of the adsorbed species, particularly with the terminating ligands, generally limits coverage to a fraction of a monolayer with each cycle. In the illustrated embodiment, roughly 0.2 Å - 0.3 Å of SiO<sub>2</sub> forms per cycle, whereas a true monolayer of SiO<sub>2</sub> is about 3 Å in height. A full monolayer forms after approximately ten cycles, where a cycle is represented by a pair of silicon source gas and oxidizing source gas pulses.

In another embodiment, a material mixture is deposited by ALD. Preferably, the mixture comprises two or more oxide species. The exemplary embodiment describes a mixture of silicon oxide (or silica) and aluminum oxide (or alumina). An exemplary process employs the primary cycle of Table I in conjunction with the secondary cycle of Table II below. In other arrangements, the skilled artisan will appreciate that other materials can be mixed together, or that other deposition processes can achieve the desired mix.

For the ALD aluminum oxide cycles, it is preferable to use temperature and pressure conditions that are close to or identical with the conditions used for the silicon oxide cycle because the pulse times are very short and it is not desirable to change these conditions between cycles.

Preferred precursor gases for aluminum deposition by ALD include aluminum halides (*e.g.*,  $\text{AlCl}_3$ ,  $\text{Al}_2\text{Cl}_6$ ,  $\text{AlBr}_3$ , and  $\text{AlI}_3$ ); alkyl aluminum compounds (*e.g.*, triethyl aluminum or  $(\text{CH}_3\text{CH}_2)_3\text{Al}$ , trimethyl aluminum or TMA or  $\text{Al}(\text{CH}_3)_3$  and others commercially available from Albemarle Corporation of USA); aluminum alkoxides (*e.g.*, aluminum isopropoxide or  $\text{Al}[\text{OCH}(\text{CH}_3)_2]_3$ , aluminum ethoxide or  $\text{Al}(\text{OC}_2\text{H}_5)_3$  and others available commercially from Strem Chemicals, Inc. of USA); aluminum beta-diketonates (*e.g.*,  $\text{Al}(\text{acac})_3$ ,  $\text{Al}(\text{thd})_3$ ,  $\text{Al}(\text{hfac})_3$  and others commercially available from Strem Chemicals, Inc. of USA); and anhydrous aluminum nitrate. The synthesis of anhydrous aluminum nitrate has been described by G. N. Shirokova, S. Ya. Zhuk and V. Ya. Rosolovskii in Russian Journal of Inorganic Chemistry, vol. 21, 1976, pp. 799-802, the disclosure of which is incorporated herein by reference. The aluminum nitrate molecule breaks into aluminum oxide when it is contacted with organic compounds, such as ethers.

In the exemplary process, the aluminum source gas comprises trimethyl aluminum mixed with inert nitrogen gas, provided at a rate of between about 50 sccm and 400 sccm, more preferably between about 100 sccm and 200 sccm. The aluminum source gas flow is maintained for between about 0.2 second and 1.0 second under the preferred reactor conditions (as noted, preferably the same as those selected for silicon oxide deposition cycles), and more preferably for between about 0.3 second and 0.6 second. A monolayer of aluminum complex chemisorbs onto the substrate or the previously formed layer thereon. The surface termination of the monolayer inhibits further reaction with the aluminum source gas.

After the aluminum complex is deposited, and excess aluminum precursor gas and reaction byproducts are purged from the chamber, an oxidizing or oxidant pulse is provided. In the illustrated embodiment, the oxidizing source gas is most preferably water vapor. The water vapor is preferably injected into the carrier gas flow or inactive gas is bubbled through liquid water, forming moist oxidizing source gas. Under the preferred temperature and pressure conditions, the water vapor pulse is maintained for between about 0.2 second and 2.0 seconds, more preferably between about 0.5 second and 1.0 second. The purge gas pulse that follows lasts for between about 0.5 second and 3.0 seconds, more preferably between about 0.5 second and 1.0 second.

In other arrangements, if ozone is used as the oxidant, it is pulsed for between about 0.5 second and 5.0 seconds, more preferably between about 1.0 second and 3.0 seconds. The purge gas pulse that follows lasts for a duration of between about 0.3 second and 4.0 seconds, more preferably between about 0.5 second and 2.0 seconds. This ALD cycle is repeated until the desired aluminum oxide thickness is reached or, as for the preferred embodiment, ALD cycles for aluminum oxide and silicon oxide are run alternately.

Table II below provides an exemplary process recipe for one cycle of aluminum oxide deposition.

TABLE II:  $\text{Al}_2\text{O}_3$

Phase	Carrier Flow (slm)	Reactant	Temperature (°C)	Pressure (Torr)	Time (sec)
aluminum	0.8	trimethyl aluminum	300	10	0.5
purge	0.8	--	300	10	0.8
oxidant	0.8	water vapor	300	10	0.8
purge	0.8	--	300	10	0.8

To achieve the desired material mixture in the illustrated embodiment, cycles of aluminum oxide deposition (such as Table II) are mixed with cycles of silicon oxide deposition (such as Table I) in an overall thin film ALD process. As noted with respect to silicon oxide, the term "aluminum oxide" or " $\text{Al}_2\text{O}_3$ " cycle is used to refer to the steps of Table II, though when mixed with other cycles it may form other materials. Table II is also referred to herein as representing a "secondary cycle." Use of the term "secondary" is for labeling purposes only, and does not implying relative importance of the material formed thereby nor the sequence in which different ALD cycles are employed. The aluminum oxide cycles and silicon oxide cycles are continued in a ratio to achieve the concentrations desired until the target thickness for the mixed material layer is reached.

Figure 4 is a gas flow diagram that shows exemplary process steps of the illustrated embodiment, illustrating two different cycles 401a and 401b, each shown with silicon oxide ALD cycles and aluminum oxide ALD cycles. Each of the cycles 401a, 401b can be considered "mullite cycles" since consecutive silicon oxide and

aluminum oxide cycles that make up the illustrated cycles 401a and 401b will generally be used in forming mullite phases. The skilled artisan will readily appreciate, however, that the ratio of silica cycles to alumina cycles to produce mullite ( $3\text{Al}_2\text{O}_3 \cdot 2\text{SiO}_2$ ) is not 1:1.

5 As has been described above, there is a continuous flow of a carrier gas 400 throughout the process. A first pulse or spike 402a of the silicon source gas is provided to form the first self-terminated silicon monolayer. After a first purge step 403, during which carrier gas continues to flow until the silicon source gas and reaction byproducts have been removed from the chamber, a first oxygen source gas pulse 404a is provided  
10 to convert the silicon layer to silicon dioxide. After a second purge 403, a pulse of aluminum source gas 406a is provided to form no more than a monolayer of aluminum in a self-terminated reaction, followed by a third purge 403. A second oxygen source gas 408a is provided to convert aluminum to alumina, followed by a fourth purge 403, thus ending the exemplary first complete cycle. The cycle can be repeated 401b in steps  
15 402b through 408b with a purge 403 at the end. The ratio of the number of aluminum source cycles to the number of silicon source cycles can be adjusted, for example as M:N, where M and N are integers from 1 to 10 (between 10:1 and 1:10), to produce the desired concentration of these materials in the final product.

#### **Preferred Trench Isolation Structure**

20 Figure 5 is a schematic illustration of an exemplary integrated circuit 200 comprising an array of active devices 202, illustrated as integrated transistors. The active devices 202 are effectively isolated from each other by trench isolation structures 204. The trench isolation structures 204 form an effective barrier against parasitic currents and capacitances that might otherwise arise between active devices 202. In  
25 addition, the space between active devices 202 and, consequently, the width of the trench isolation structure 204 therein, should be minimized to increase the density of active devices 202 within an integrated circuit 200.

In the preferred embodiment, the width of the trench 204 is preferably less than about 100 nm, more preferably less than about 70 nm. The depth of the isolation trench  
30 can vary according to the isolation requirements of the specific integrated circuit. Accordingly, it can vary from as shallow as 0.2  $\mu\text{m}$  to as deep as 1.0  $\mu\text{m}$ . In the



preferred embodiment, the aspect ratio of the trench, which is the ratio of the depth to the width, is preferably greater than 3:1, more preferably, greater than about 5:1.

5        Figures 6 through 9 are cross-section drawings that illustrate the sequence for forming a trench isolation structure of the preferred embodiment. In Figure 6, trench 204 is formed within a semiconductor substrate 100. Conventional photolithography techniques are employed to define the trench position and shield the rest of the wafer from the subsequent etch. Anisotropic etching, such as reactive ion etching, is preferred for producing a trench with vertical sidewalls.

10        Figure 6 illustrates an insulating thin film 206a that has been built up by a series of monolayers through the ALD process as described above. The thin film 206a can comprise any suitable insulating material, including silicon dioxide. More preferably, however, the material is selected to have a coefficient of thermal expansion (CTE) that matches that of the surrounding substrate. Thin film 206a represents a plurality of ALD cycles, producing a film that is highly conformal to all surfaces of the trench 204, including the vertical sidewalls 208 and bottom 210, and also to the top surface 212 of the substrate 100.

15        Figure 7 illustrates an isolation layer 206b at a subsequent deposition stage. The isolation layer 206b is a conformal continuation of the isolation layer 206a. It should be generally understood that the dashed line between layers 206a and 206b is arbitrary. Layer 206a indicates simply a "snapshot" of the deposition at some earlier time,  $t_a$ . It is not actually a discrete layer. Layer 206b indicates the deposition that has occurred between time  $t_a$  and a later time  $t_b$ . The isolation layer 206b consists of a series of monolayers that build up on top of the isolation layer 206a.

20        Figure 8 illustrates an isolation layer 206 at a stage when the trench has been filled. The isolation layer 206c is a conformal continuation of the isolation layer 206b and is arbitrarily designated by a dashed line. At the final stage, the preferred layer 206c seals and fills the trench 204 completely, advantageously forming minimal or no discernible seam where the layers from opposing sides of the trench unite at the center. This is an advantage over chemical vapor deposition processing wherein a seam can form at the center of the trench and can even contain voids, particularly when subjected to thermal cycling.

25       

30

As shown in Figure 9, trench isolation structure 214 is preferably planarized, such as by chemical mechanical polishing (CMP) or a reactive ion etchback process that does not disturb the isolation layer 206 inside the trench 204. Figure 9 illustrates the trench isolation structure 214 after the isolation material has been removed from the top surface 212 of the substrate 100 by planarization. The trench isolation structure 214 has a planar surface 216 at its top opening that is level with the top surface 212 of the substrate. This planarity is advantageous for subsequent conformality of deposited layers and photolithographic processes.

In another embodiment, a trench isolation structure is filled with ALD material as described above, except that two different materials are deposited, for example, as outlined in the gas flow diagram of Figure 4. These materials may not be in equilibrium at the processing temperatures of the semiconductor wafer, in which case, they undergo a phase transformation to an equilibrium two-phase mixture. In yet another embodiment, more than two different materials are deposited. Deposited materials can be chosen to effect multi-phase mixtures with desirable properties, such as coefficient of thermal expansion (CTE) or dielectric constant.

In an exemplary embodiment, the trench is filled with an insulating material selected to match the CTE of the surrounding substrate. In the illustrated example, the isolation material comprises a mixture of silica (nominally  $\text{SiO}_2$ ) and alumina (nominally  $\text{Al}_2\text{O}_3$ ). These two materials do not exist in equilibrium, especially at the temperatures to which the wafer is exposed during integrated circuit fabrication. Also, with the two materials deposited by selected ratios of alumina to silica cycles in a continuous ALD process, as described above, there is no significant diffusion kinetics barrier to overcome in reaching equilibrium, so a transformation to the equilibrium structure can occur. Optionally, after deposition, an annealing step can be used to densify and remove water from the insulating mixture.

Figure 10 shows the silica-mullite phase diagram from W. E. Lee and W. M. Rainforth, Ceramic Microstructures, Chapam and Hall, p. 297 (1994). At temperatures less than about  $1600^\circ\text{C}$ , the equilibrium result of a mixture containing less than about 70 weight percent alumina with silica is a two-phase mixture of silica ( $\text{SiO}_2$ ) and mullite ( $3\text{Al}_2\text{O}_3 \cdot 2\text{SiO}_2$ ). Silica and mullite are continuously miscible in this composition range.

To a first approximation, their coefficient of thermal expansion (CTE) scales linearly. If a dielectric mixture with a CTE close to the CTE of silicon is used to fill the silicon trench, it is able to withstand large changes in temperature without stress at the dielectric/silicon interface.

5           Line 300 in Figure 11 shows the coefficient of thermal expansion as a function of weight percent of  $\text{Al}_2\text{O}_3$  in a mixture with  $\text{SiO}_2$ . The CTE of silicon is shown by the horizontal line 310 at  $2.3 \times 10^{-6} / \text{K}$ . The slope of mixture line 300 is not very steep near its intersection 320 with silicon line 310. At the intersection 320, a composition of about 30 wt % (weight percent)  $\text{Al}_2\text{O}_3$  and about 70 wt %  $\text{SiO}_2$ , which is a two-phase mixture of about 40 wt % mullite and about 60 wt % silica at equilibrium, has the same  
10           coefficient of thermal expansion (CTE) as silicon, i.e.,  $2.5 \times 10^{-6} / \text{K}$ .

          Preferably, the mixture composition is such that the CTE falls within about plus or minus 20% of the CTE of silicon, i.e., in the range from about  $2.0 \times 10^{-6} / \text{K}$  at point 330 to about  $3.0 \times 10^{-6} / \text{K}$  at point 340. This corresponds to an  $\text{Al}_2\text{O}_3$  concentration  
15           between about 23 wt % and about 37 wt % in the mixture, and, after a phase transformation to thermal equilibrium, the two-phase mixture ranges from about 25 wt % mullite/75 wt % silica to about 50 wt % mullite/50 wt % silica. More preferably, the mixture's CTE falls within about plus or minus 10% of the CTE of the surrounding material, or between about  $2.25 \times 10^{-6} / \text{K}$  and  $2.75 \times 10^{-6} / \text{K}$ . This corresponds to an  
20            $\text{Al}_2\text{O}_3$  concentration in the mixture between about 26 wt % and about 34 wt %, and, after a phase transformation to thermal equilibrium, the two-phase mixture ranges from about 35 wt % mullite/65 wt % silica to about 40 wt % mullite/60 wt % silica.

          Preferably, to achieve a CTE within about 20% of silicon's CTE, ALD can be conducted with a ratio of silica cycles to alumina cycles of between about 20:1 and  
25           1:10. More preferably, a CTE within about 10% of silicon's CTE can be achieved with a ratio of silica ("primary") cycles to alumina ("secondary") cycles of between about 10:1 and 3:1

          Accordingly, there are several advantages in the described invention. For example, ALD material conformally fills a narrow opening, particularly a trench isolation  
30           structure, which advantageously allows for increased the packing density of active devices in an integrated circuit. Additionally, due to ALD conformality, trench sidewalls can be

formed vertically to increase the packing density of integrated circuit devices. Furthermore, ALD technology permits precise control over thin layer dimensions, thus allowing for custom tailoring of trench isolation devices with very narrow and deep openings, as needed for device isolation requirements.

5 By understanding the phase transformations that occur with each combination of materials deposited, a wider variety of material properties can be engineered than from the as-deposited materials alone. Concentration ratios can be carefully chosen to effect transformation to specific two-phase mixtures that have desirable properties, such as linear coefficient of thermal expansion or dielectric constant, for trench isolation structures.

10 The skilled artisan will readily appreciate, in view of the disclosure herein, that other methods can be employed to achieve the desired CTE matching of filler material with surrounding trench. ALD, however, advantageously provides excellent process control to produce near-perfect step coverage (preferably greater than 95%, more preferably greater than 98%) over high aspect ratio trenches. Furthermore, the ALD process can also be readily tuned by adjusting ratios of reactant pulses or cycles, as described herein, to attain the desired composition.

15 Furthermore, though described with particular materials and in the context of shallow trench isolation, the skilled artisan will readily find application for the principles and advantages described herein for matching CTE of other materials and structures in the course of integrated circuit fabrication. Another example is the filling of a contact via with conductive material having a CTE matched to that of surrounding insulating material. As newer materials are developed for advanced fabrication of ever-faster circuits, the flexibility of the processes described herein can be readily adapted to address CTE matching for various materials in integrated circuits.

20 Thus, although the foregoing invention has been described in terms of certain preferred embodiments, other embodiments will become apparent to those of ordinary skill in the art in view of the disclosure herein. Accordingly, the present invention is not intended to be limited by the recitation of preferred embodiments, but is intended to be defined solely by reference to the appended claims.